

## Conference Abstract

# Parasitism in Eukaryota - Reconstruction of Ancestral and Unavailable Extant States

Lydia R. E. C. Buntrock<sup>‡</sup>, Bernhard Y. Renard<sup>§</sup>, Emanuel Heitlinger<sup>‡,|</sup>

<sup>‡</sup> Institute for Biology, Molecular Parasitology, Humboldt University, Berlin, Germany

<sup>§</sup> Bioinformatics Unit (MF1), Department for Methods Development and Research Infrastructure, Robert Koch Institute, Berlin, Germany

<sup>|</sup> Research Group Ecology and Evolution of Molecular Parasite Host Interactions, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany

Corresponding author: Lydia R. E. C. Buntrock ([lydia.buntrock@irallia.de](mailto:lydia.buntrock@irallia.de)), Bernhard Y. Renard ([renardb@rki.de](mailto:renardb@rki.de)), Emanuel Heitlinger ([emanuel.heitlinger@hu-berlin.de](mailto:emanuel.heitlinger@hu-berlin.de))

Received: 12 Apr 2018 | Published: 21 May 2018

Citation: Buntrock L, Renard B, Heitlinger E (2018) Parasitism in Eukaryota - Reconstruction of Ancestral and Unavailable Extant States. Biodiversity Information Science and Standards 2: e25739.

<https://doi.org/10.3897/biss.2.25739>

## Abstract

Parasitism can be defined as an interaction between species in which one of the interaction partners, the parasite, lives in or on the other, the host. The parasite draws food from its host and harms it in the process. According to some estimates, over 40% of all eukaryotes are parasites. Nevertheless, it is difficult to obtain information about a particular taxon is a parasite computationally making it difficult to query large sets of taxa.

Here we test to what extend it is possible to use the Open Tree of Life (OTL), a synthesis of phylogenetic trees on a backbone taxonomy (resulting in unresolved nodes), to expand available information via phylogenetic trait prediction. We use the Global Biotic Interactions (GloBI) database to categorise 25,992 and 34,879 species as parasites and free-living, respectively, and predict states for over ~2.3 million (97.34%) leaf nodes without state information.

We estimate the accuracy of our maximum parsimony based predictions using cross-validation and simulation at roughly 60-80% overall, but strongly varying between clades. The cross-validation resulted in an accuracy of 98.17% which is explained by the fact that

the data are not uniformly distributed. We describe this variation across taxa as associated with available state and topology information. We compare our results with several smaller scale studies, which used manual expert curation and conclude that computationally inferred state changes largely agree in number and placement with those. In clades in which available state information is biased (mostly towards parasites, e.g. in Nematodes) phylogenetic prediction is bound to provide results contradicting conventional wisdom.

This represents, to our knowledge, the first comprehensive computational reconstruction of the emergence of parasitism in eukaryotes. We argue that such an approach is necessary to allow further incorporation of parasitism as an important trait in species interaction databases and in individual studies on eukaryotes, e.g. in the microbiome.

## Keywords

Ancestral state reconstruction, parasites, maximum parsimony, Castor, Global biotic interactions (GloBI), Open tree of life (OTL)

## Presenting author

Lydia Bunrock